Rec'd PCT/PTO 08 DEC 2004

MULTIMEDIA SERVER WITH SIMPLE ADAPTATION TO DYNAMIC NETWORK LOSS CONDITIONS

FIELD OF THE INVENTION

[0001] This invention relates towards the field of transmitting prioritized data based on network conditions.

BACKGROUND OF THE INVENTION

[0002] With the development of communications networks (network fabric) such as the Internet and the wide acceptance of broadband connections, there is a demand by consumers for video and audio services (for example, television programs, movies, video conferencing, radio programming) that can be selected and delivered on demand through a communication network. Video services, referred to as media objects or streaming audio/video, often suffer from quality issues due to the bandwidth constraints and the bursty nature of communications networks generally used for streaming media delivery. The design of a streaming media delivery system therefore must consider codecs (encoder/decoder programs) used for delivering media objects, quality of service (QoS) issues in presenting delivered media objects, and the transport of information over communications networks used to deliver media objects, such as audio and video data delivered in a signal.

[0003] Codecs are typically implemented through a combination of software and hardware. This system is used for encoding data representing a media object at a transmission end of a communications network and for decoding data at a receiver end of the communications network. Design considerations for codecs include such issues as bandwidth scalability over a network, computational complexity of encoding/decoding data, resilience to network losses (loss of data), and encoder/decoder latencies for transmitting data representing media streams. Commonly used codecs utilizing both Discrete Cosine Transformation (DCT) (e.g., H.263+) and non-DCT techniques (e.g., wavelets and fractals) are examples of codecs that consider these above detailed issues. Codecs are also used to compress and decompress data because of the limited bandwidth available through a communications network.

[0004] Quality of service issues relate to the delivery of audio and video information and the overall experience for a user watching a media stream. Media objects are delivered through a communications network, such as the Internet, in discrete units known as packets. These units of information, typically transmitted in sequential order, are sent via the Internet through nodes commonly known as servers and routers. It is therefore possible that two sequentially transmitted packets arrive at a destination device at different times because the packets may take different paths through the Internet. Consequentially, a QoS problem known as dispersion could result where a packet transmitted later in time may be processed and displayed by a destination device before an earlier transmitted packet, leading to discontinuity of displayed events. Similarly, it is possible for packets to be lost when being transmitted. A destination device typically performs an error concealment technique to hide the loss of data. Methods of ensuring QoS over a network such as over-allocating the number of transmitted packets or improving quality of a network under a load state may be used, but these methods introduce additional overhead requirements affecting communication network performance.

[0005] Communication networks control the transfer of data packets by the use of a schema known as a transport protocol. Transmission Control Protocol (TCP), described in Internet Engineering Task Force (IETF) Request For Comments (RFC) 793, is a well-known transport protocol that controls the flow of information throughout a communications network. A transport protocol attempts to stabilize a communications network by maintaining parameters such as flow control, error control, and the time-organized delivery of data packets. These types of controls are administered through the use of commands that exist in a header of a packet or separately from packets transmitted between devices through the communications network. This control information works well for a communications network that operates in a "synchronous" manner where the transmission of data packets tends to be orderly.

[0006] Other types of media objects, in the form of streamed data, tend to be delivered or generated asynchronous by where the flow of packets may not be consistent. These packets are transmitted and received at different times, hence asynchronously, where received packets are reconstituted in view of data in the headers of such packets. The transmission of asynchronous packets suffers when

network conditions drastically reduce the transmission (or receipt) of packets, resulting in network loss of service, degradation, or other conditions requiring a transmission to time out.

[0007] One way of reducing the amount of errors in the transmission of a data uses a technique called forward error coding (FEC) where some data is repeated in a data stream. By using FEC, other methods of error correction such as error concealment, flow control, and the like are not required for a user to acquire successfully a media object transmitted in a data stream. FEC however requires that the transmitter of data stream take into account network conditions that lead to a corruption or loss of data packets impacting an encoder that encodes data on the fly.

SUMMARY OF THE INVENTION

[0008] A method for transmitting prioritized data encoded by a Forward Error Coding operation is disclosed. A media object is separated into different classes of data, forming a base layer and at least one enhancement layer of information, with each layer having associated parity data. Data of the separated media object, formed of classified data, is later encoded and stored, whereby information of the base layer is assigned a higher priority for transmission than enhancement layer data. Such priority classifications are used when a server transmits a composition of classified data over a network fabric, as prioritized data.

BRIEF DESCRIPTION OF THE DRAWINGS

[0009] FIG. 1 is a diagram of a system illustrating the prioritization, encoding, and transmission of a media object, according to an illustrative embodiment of the invention.

[0010] FIG. 2 is a block diagram of a method for generating and transmitting classified data representing a media object as prioritized data, according to an illustrative embodiment of the invention.

[0011] FIG. 3 is a block diagram of method decoding prioritized data representing a media object, according to an illustrative embodiment of the invention.

DETAILED DESCRIPTION OF THE INVENTION

[0012] As used herein, multimedia related data that is encoded and is later transmitted represents a media object. The terms information and data are also used synonymously throughout the text of the invention as to describe pre or post encoded audio/video data. The term media object includes audio, video, textual, multimedia data files, and streaming media files. Multimedia files comprise any combination of text, image, video, and audio data. Streaming media comprises audio, video, multimedia, textual, and interactive data files that are delivered to a user's device via the Internet or other communications network environment and begin to play on the user's computer/ device before delivery of the entire file is completed. One advantage of streaming media is that streaming media files begin to play before the entire file is downloaded, saving users the long wait typically associated with downloading the entire file. Digitally recorded music, movies, trailers, news reports, radio broadcasts and live events have all contributed to an increase in streaming content on the Web. In addition, the reduction in cost of communications networks through the use of high-bandwidth connections such as cable, DSL, T1 lines and wireless networks (e.g., 2.5G or 3G based cellular networks) are providing Internet users with speedier access to streaming media content from news organizations, Hollywood studios, independent producers, record labels and even home users themselves.

[0013] The preferred embodiment of the invention makes use of a subset of FEC techniques known as forward erasure correction (FXC) where the content of a media object is pre-encoded into separate partitions. Using techniques known in the art, a media object is encoded into different classes of data, referred to as classified data. Each class of data represents a different layer of information (i.e., a base and enhancement layers) where the base layer represents data crucial for rendering a media object and the enhancement layers being data that is less critical but important for adding detail to a rendered media object.

[0014] The classified data is further refined by using systematic FXC codes, such as Reed Solomon (RS) codes, as to create parity data that is transmitted with the data representing base and enhancement layers of an encoded media object. Specifically, RS is used to produce erasure codes of various strengths whereby overhead rates for communication data can be generated using a RS code with

different (n, k) parameters; n equal to the total amount of data to be transmitted (encoded layer data with parity data) and k equal to the amount of encoded data.

[0015] When used for erasure correction, an RS code can correct up to h=n-k erasures (or the amount of data missing from a transmitted data stream). If the exemplary system uses a Galois Field with 8 bit symbols as the basis of transmitted data, the maximum value of n is calculated q=p^r (q=maximum value of n, p=amount of data states, r=number of items with data states). Hence, for an 8 bit symbols, p=2 (a bit having two states) and r=8 (number of bits), the maximum value of n is 255.

[0016] Shorter length FXCs can be used by only computing and transmitting as many parity bits that are as desired or needed. Once a maximum n is calculated, a smaller RS(n', k) may be derived from a RS(n, k) code where n'<n, which is modified depending on the desired erasure protection strength (see, L. Rizzo, "Effective Erasure Codes for Reliable Computer Communications Protocols", Computer Communication Review, 27(2): pgs. 24-36, April 1997) The calculated parity bits for encoded data may change in accordance with network conditions or encoder performance.

[0017] As an example of encoding a byte based code based on a 2^8 Galois Field, a maximum value of n=255 is calculated. A RS(n',k) code is selected, where the Reed Solomon code is based on an RS(255, k), and n'-k parity bytes are encoded. As the value of n' increases, the original parity bytes encoded (n'-k) are not changed. That is, for a Reed Solomon code for a RS(11, 10) based on a RS(255, 10), the 11th parity byte has the same value as the 11th parity byte in an RS(12, 10) code. It is to be noted that the principles of the present invention may be modified to accommodate different values of n, n', p, r, and k depending on the needs of an encoding/transmitting system.

[0018] Preferably, RS coding of data is interleaved across packets or frames. That is, entire packets or frames will be made up of either information or parity data. These packets, in order to simplify the process of identifying missing packets, may be identified by information in the packet headers. Hence, a media object requester would be able to identify missing packets if the packet headers are sequentially generated, and there is a gap in the numeric sequence. Real Time Transport Protocol (RTP) is one transport mechanism used for generating sequential

packet headers, although other transport protocols may be selected in accordance with the principles of the present invention.

[0019] Additionally, different levels of channel loss protection are achieved by grouping parity packets into several multicast groups. A client receiving such data can adjust the level of channel loss protection by joining (or leaving) as many multicast groups as needed, hence the client may adapt to the loss of data by increasing the channel bandwidth by joining more multicast groups, as needed. This technique of multicasting is described because the source-encoding rate of a FXC encoder is typically is not adjusted in the case where content is pre-encoded and stored on a storage device, for an exemplary embodiment of the present invention.

[0020] When encoding a media object separated into different classes of data layers, it is desirable to offer a higher FXC strength for base layer data and a lower FXC strength for enhancement layer data is accomplished by using scalable video compression with unequal error protection. For an exemplary embodiment of the present invention, a media object is separated into two layers of classified data: base layer information (Bi) and enhancement layer information (Ei). Accordingly, the base layer has parity data (Bp) and the enhancement layer has parity data (Ep); each of layer and parity data are afforded their own data types. Bi and Bp is data that is more important than Ei and Ep data, because Bi and Bp data is more critical for rendering media object than Ei and Ep data. It should be noted that the principles of the present invention apply where a media object is prioritized into as many layers as needed, for example, one base layer and multiple enhancement layers.

[0021] An exemplary embodiment of the invention, shown as encoding system 100 in FIG. 1, presents scalable video encoder 110 that creates compressed bit streams from a media object being encoded. Scalable video encoder 110 may be implemented in software, hardware, or in a combination of both. The media object is divided into separate layers of classified data as described above, where the data once separated, is placed in a bitstream corresponding to a priority assigned to each layer and packed into packets for network transmission via network fabric 160, such as a communications network or the Internet. Preferably, each layer is FXC encoded, using a systematic FEC encoder 115, 120 across packets for protection against network packet loss. The priority of each layer of classified data is

associated with the importance of transmitted data eventually used to render a media object.

[0022] More specifically in this exemplary embodiment, scalable video encoder 110 separates the media object into two layers, representing a base layer and an enhancement layer. Data representing the base layer is inputted into FEC encoder 115 where Bi information is generated via a FXC encoding process. This generated data is stored as pre-encoded data in Bi storage 125. FEC encoder 115 also creates Bp data that is stored in Bp storage 130 when generating Bi information.

[0023] Similarly, data representing the enhancement layer is inputted into FEC encoder 120 where Ei information is generated via a FXC encoding process. This generated data is stored as pre-encoded data in Ei storage 135. FEC encoder 115 also creates Ep data that is stored in Ep storage 140 when generating Ei information. Different strength FXC codes can be used for the base and enhancement layers, depending on network and system requirements. Preferably, when adjusting the FXC strength of transmitted RS codes, an indication the contents of data packets is transmitted, either in data packet headers or as separate side information.

[0024] When a request is made for a media object via network fabric 160, multimedia server 150 preferably determines the available bandwidth and expected (or real time) network loss conditions that effect the requester of the media object. This type of determination may be made based on a user profile, information communicated in the request for a media object, historical network conditions, network service reporting information (such as Real time Transport Control Protocol (RTCP) reports obtained during the transmission of data), and the like. Optionally, multimedia server 150 determines the type of network path to be used to deliver the pre-encoded media object to estimate a possible network loss. For example, multimedia server 150 expects a higher loss rate of data when a wireless connection is used versus a landline or broadband connection to communicate data.

[0025] Multimedia server 150, in response to the determination of network conditions, selects Bi, Bp, Ei, and Ep data from their associated storage areas based on the level of priority assigned to the selected data. This priority level is related to the importance of the data as used to render a media object. Hence, base layer data is considered more important and is more likely to be transmitted than enhancement

layer data during periods of network congestion. After selecting classes of data to be transmitted, multimedia server 150 creates a composition of classified data by prioritizing and formatting such selected data. This composition of classified data, known as prioritized data, reflects multimedia encoder 150 adjusting the classes of data transmitted in view of network conditions, where a minimum level of base layer information is required to render a media object. As network conditions improve, the composition of classified data includes more enhancement layer information and associated parity information.

[0026] Multimedia server 150 transmits data packets of prioritized data over network fabric 160. Specifically, multimedia server 150 seeks to optimize the playback quality of multimedia data received by a requestor of a media object by adjusting the composition of Bi, Bp, Ei, and Ep transmitted in accordance with their respective priority classifications. For example, if no loss of data is expected from a network, multimedia server transmits all of the Bi and Ei information in data packets. Bp and Ep data is transmitted as space/bandwidth allows, preferably with more Bp data being transmitted than Ep data.

[0027] When there is an expected level of network loss, multimedia server 150 replaces Ep data with Bp data in the composition forming prioritized data. With very high levels of expected network loss, multimedia server replaces an amount of Ei information transmitted with Bp data because a requested media object will not be capable of being rendered without a baseline of Bi information that is received or recovered using Bp data. It is to be noted that there may be a limit to the bandwidth available to a media object requester due to physical or pre-set bandwidth limits of a network.

[0028] In an optional embodiment of the present invention, multimedia server 150 attempts to optimize the delivery of a media object to a requestor by determining the amount of expected network loss, as explained above. Assuming that the bandwidth to a requestor is fixed, multimedia server 150 transmits a composition of Bi information and an amount of Bp data necessary to achieve a corrected error rate, in response to the expected network loss. If there is any available bandwidth after the transmission of Bi and Bp data, multimedia server 150 fills the space first with Ei and then Ep data. The tradeoff between transmitting Bp versus Ei or Ep depends on many factors such as the expected range of network loss

conditions, effectiveness of the scalable encoding, viewer preferences, nodes in a network, and the like.

[0029] Preferably, multimedia server 150 will use high strength FXC codes when transmitting Bi, Bp, Ei, and Ep data representing an encoded media object. By using system 100, stored FXC codes will not need to be recomputed each time expected network conditions change for a new media object requestor.

[0030] In the operation of encoding system 100, a temporal encoding technique is preferred over spatial, Signal Noise Ratio (SNR), or simple data partitioning encoding techniques because temporal based processes do not suffer from the problem of "drift". Specifically, when decoding an media object that has been prioritized and separated into layers, periods of drift occur when decoding base and enhancement layer data after exclusively decoding base layer data. The reconstructed media object (especially video) rendered from the base and enhancement layer data will continue to appear as if it were being rendered during the time of just base layer data. This drift effect is minimized if base and enhancement layers were exclusively used for decoding a media object.

[0031] The problem of drift is eliminated when temporally encoded video based media objects place bidirectional "B" coded pictures in the enhancement layer, and "I" and "P" frames are placed in the base layer. Preferably, the B coded pictures in the enhancement layer are not used to predict other pictures. Hence, when media server 150 transmits Bp data instead of Ei information, a media object requestor's video frame rate is reduced, but the per frame video quality is not reduced if the FXC code strength is sufficient to correct all network loss.

[0032] During periods of network disruption, a media object requestor would use correctly received Ei information to increase the frame rate of video, which is be greater than the frame rate of video using only base layer data. When network conditions improve, more Ei information is transmitted, and the frame rate of the video will likewise improve, with the quality of rendered video. Optionally, the media object requestor (or decoder of the media object requestor) may request that the composition of transmitted Bi, Bp, Ei, and Ep data as priority data be changed in accordance with network conditions. Multimedia server 150 implements this request.

[0033] Ideally, Bi, Bp, Ei, and Ep data are packed into data packets, where fixed sizes of data packets are used. Multimedia server 150 is able to swap entire

data packets during transmission, as to maintain a constant data transmission rate. A drawback to this technique however prevents a correspondence between video frames or slices and data packets, as suggested in IETF RFC 2250 and RFC 2190. An alternative embodiment of the present invention is supported where the data packets do correspond to video frames or slices, which depends upon the technique selected for packing and processing data packets.

[0034] FIG. 2 represents block diagram of a method 200 for the transmission of prioritized data representing a media object by multimedia server 150, in accordance with an exemplary embodiment of the present invention. In step 210, scalable video encoder 110 and FEC encoders 115 and 120 encode a media object into levels of classified data. Specifically, scalable encoder 110 separates a media object into several classes of data, denoted as separate layers, with each layer corresponding to the importance of data used for rendering a media object. The layers of data form a base layer and at least one enhancement layer(s) of information. The separated layers of classified data are relayed to FEC encoders 115 and 120 for FXC encoding. During the encoding process, parity data associated to each layer is generated and is later stored in step 220. Importantly, the generated information and parity data corresponding to each layer is stored in their respective storage areas, such as base layer information being stored in Bi storage 125 and the associated priority information being stored in Bp storage 130. Optionally, there are as many storage areas as there are layers of classified data.

[0035] Multimedia server 150, in response to a request for a media object, prioritizes a composition of classified data into prioritized data and transmits such data in response to network conditions in step 230. The prioritization of classified data is determined by the level of priority assigned to each layer of classified data. Multimedia server 150 forms the composition of classified data, as prioritized data, in view of network conditions. When network conditions result in the loss of data, data with a higher priority level is more likely to be transmitted than data with a lower priority level. Conversely, data with a lower priority is more likely to be transmitted when network conditions result in fewer data packets being loss.

[0036] The determination of network conditions, as described above, may either be expected or real-time network conditions. Accordingly, multimedia server 150 retrieves data from storage 125, 130, 135, and 140 in accordance with network

conditions. If a network encounters many problems, more Bi, and Bp data is retrieved and transmitted over network fabric 160, versus periods of no network problems where more Ei and Ep data is transmitted.

[0037] Multimedia server 150 adjusts the composition of classified data, forming prioritized data, in response to a change in network conditions in step 240. If network conditions improve, multimedia server 150 will transmit more enhancement layer related information (Ei, Ep). If network conditions worsen during a transmission, multimedia server 150 will replace enhancement layer associated data with more base layer associated data (Bi, Bp). This process may be repeated between steps 230 and 240 as network conditions change frequently.

[0038] FIG. 3 presents a block diagram of method 300 for an exemplary embodiment of a decoder decoding prioritized data operating in accordance with the principles of the present invention. Specially, in step 310 a media object requester makes a request for a media object via network fabric 160. Multimedia server 150 preferably receives this request, where the present network conditions of the requester are communicated with the request.

[0039] In step 320, a decoder used by the media object requester begins to process received prioritized data, wherein such data preferably has at least Bi information. The decoder uses prioritized data formed of a composition of classified data to render a media object as audio, video, or a combination of both. If the decoder receives more Ei data, a decoder renders a media object at a higher level of quality than possible with just Bi information. The receipt of parity data related to either the base layer or enhancement layer(s) assists in the generation of missing Bi or Ei information if network conditions result in the loss of transmitted data.

[0040] In an optional embodiment of the invention, a decoder uses FXC decoding if data was lost during the receipt of data packets representing a media object. Specifically, the decoder may not receive all of the transmitted data representing either Bi or Ei information. By using FXC decoding, the decoder generates missing Bi information from received Bp data and missing Ei information from received Ep data.

[0041] The decoder, in step 330, requests that the composition of classified data transmitted as prioritized data change, because network conditions are different. Specifically, the decoder either requests that enhancement layer information be

replaced by base layer parity data, for degrading network conditions, or for more enhancement layer or parity data for improving network conditions. The mechanics of the decoder of the media object requestor is similar to the inverse of the operation of scalable video encoder 110.

[0042] The present invention may be embodied in the form of computer-implemented processes and apparatus for practicing those processes. The present invention may also be embodied in the form of computer program code embodied in tangible media, such as floppy diskettes, read only memories (ROMs), CD-ROMs, hard drives, high density disk, or any other computer-readable storage medium, wherein, when the computer program code is loaded into and executed by a computer, the computer becomes an apparatus for practicing the invention. The present invention may also be embodied in the form of computer program code, for example, whether stored in a storage medium, loaded into and/or executed by a computer, or transmitted over some transmission medium, such as over electrical wiring or cabling, through fiber optics, or via electromagnetic radiation, wherein, when the computer program code is loaded into and executed by a computer, the computer becomes an apparatus for practicing the invention. When implemented on a general-purpose processor, the computer program code segments configure the processor to create specific logic circuits.